PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (http://bmjopen.bmj.com/site/about/resources/checklist.pdf) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Trends in the incidence of testing for vitamin D deficiency in
	primary care in the United Kingdom: a retrospective analysis of
	The Health Improvement Network (THIN), 2005 to 2015
AUTHORS	Crowe, Francesca; Jolly, Kate; MacArthur, Christine; Manaseki-
	Holland, Semira; Gittoes, N; Hewison, M; Scragg, Robert;
	Nirantharakumar, Krishnarajah

VERSION 1 - REVIEW

REVIEWER	Emre Basatemur Institute of Child Health University College London UK
REVIEW RETURNED	23-Dec-2018

GENERAL COMMENTS	This manuscript describes analysis of a UK primary care research database (THIN), investigating trends in testing of vitamin D levels over time among adults in primary care and socio-demographic factors associated with rates of testing. Similar work has been previously published reporting rates of vitamin D testing and associated healthcare costs in children in primary care, using THIN (European Journal of Pediatrics, 2017;176(10):1405-1409). The current manuscript describes a similar time trend in testing to that which has been reported in children, with a marked increase in testing over the last decade.
	The analysis has generally been conducted appropriately, and the paper is clear and well written. Overall the work is a useful addition to the literature. However, there are a number of points which I believe should be addressed prior to publication. These are detailed below. Abstract:
	1. The wording for the reporting of 25-OH-D deficiency prevalence in the abstract is currently misleading. In the conclusion the authors state that 'One-third of the UK population were vitamin D deficient'. However, the study is not an analysis of vitamin D levels in a random sample of the population, and the prevalence of deficiency among those tested is unlikely to be representative of that in the population as a whole. This should be re-worded to state 'One third of UK adults who had a vitamin D test performed in primary care were found to be vitamin D deficient'.
	in primary care were found to be vitamin D deficient'. Methods:

- 1. Why did the authors include 1,25-dihydroxyvitamin D tests in the analysis? This is not a measure of vitamin D status nor a test for vitamin D deficiency. It is a specialist test indicated in patients with renal disease or parathyroid dysfunction. It is very different from 25-OH-D tests used to assess vitamin D status, and it should not be combined with 25-OH-D tests as a single outcome. I would recommend removing 1,25-dihydroxyvitamin D tests from the study outcome, as it is not relevant to the aims of the study (time trends in testing for vitamin D deficiency).
- 2. I find the methods used for handling missing data for ethnicity and SES slightly confusing. For the main analysis, the authors have handled missing data by introducing a separate category indicating missingness. However, they have then not included this category when reporting the results in Table 2 (instead giving the complete results for this model in Supplementary Table 6). Table 2 should include all categories for each variable included in the model, including the one for missing. It is misleading to exclude a category for a variable when reporting the results. Table 2 should be updated to include the missing category, then Supplementary Table 6 removed.

As a sensitivity analysis, the authors performed the analysis using complete cases analysis. However, both are simple methods of dealing with missing data that are only valid if data are missing completely at random, and the results are unlikely to differ much between methods. If the authors wanted to undertake a more sophisticated sensitivity analysis, then they could go down the multiple imputation route. If not, then my personal preference would be for the authors to just present the main results using either the 'Indicator variable for missingness' method or the 'complete cases' method, and leave it at that (without reporting the other of these two simple methods as a "sensitivity analysis").

- 3. Why did the authors choose the cut-off of <30 nmol/l to represent vitamin D deficiency? This should be explained. The National Osteoporosis Society guidelines (which they refer to) and NICE guidelines for management of vitamin D deficiency in adults both use a cut-off of <25 nmol/l to recommend treatment, so I am surprised that the authors chose <30. It would also be helpful if the authors included a comment in the discussion about the lack of consensus internationally regarding the choice of 25-OH-D threshold to indicate deficiency.
- 4. Why did the authors exclude patients with 25-OH-D levels <5 or >200 in the analysis of prevalence of vitamin d deficiency among the tested individuals? Did they decide that these were biologically implausible 'outlier' values? I would argue that these values should not be treated as outliers, as they are clinically plausible. It is not unusual to see results of <5 in deficient individuals, and also high levels >200 can be seen in individuals taking high dose supplementation. Removing these test values will bias the results. Perhaps the authors chose to remove them because a result of "<5" or ">200" does not represent an integer value that can be fitted into a mathematical model. In this case, I would suggest a preferable solution would be to replace these results with a value of 4 or 201 respectively rather than to ignore them. Simply excluding the extreme (but plausible) test results will lead to results that are mis-representative of the true distribution of deficiency.

Results:

1. Minor point. The title for Supplementary Table 1 is missing.

Conclusions:

- 1. The authors state that the results of this paper cannot be directly compared to the published work investigating vitamin D testing and diagnosis in children in UK primary care. However, they are perhaps unaware of a paper which has reported rates of vitamin D testing among children in UK primary care (European Journal of Pediatrics, 2017;176(10):1405-1409, see Supplementary Table 1 in this paper). This paper reports very similar time trends in vitamin D testing among children. It would be helpful if the authors compare their work to this existing paediatric literature.
- 2. The authors hypothesise that the decrease in percentage of tested patients with deficient results over time may be due to changes in laboratory assay methods. However, I would argue that a much more likely explanation for this observation is that vitamin D testing is being performed less selectively (indicated by the large increase in testing rates over time).
- 3. The cost analysis on page 13 is very misleading as only one test per individual is included, and therefore represents an underestimation of true costs. If the authors wish to make a meaningful estimate for primary care expenditure on vitamin D tests in adults, they need to include all tests (including repeat tests in the same individual). Otherwise it is better to leave this out altogether. See European Journal of Pediatrics, 2017;176(10):1405-1409 for an analysis UK primary care expenditure on vitamin D tests in children, where repeat tests within individuals were included.

Article Summary:

1. The authors state that incidence rates for vitamin D testing in UK primary care have not previously been examined. This is not entirely accurate, as it has been examined in children (European Journal of Pediatrics, 2017;176(10):1405-1409). The authors should qualify that incidence rates for vitamin D testing in UK primary care have not previously been examined in adults

REVIEWER	Pascal CAILLET
	Nantes' University Hospital
REVIEW RETURNED	27-Dec-2018

GENERAL COMMENTS	The manuscript describe the results of an open cohort analysis, focusing firstly on establishing the incidence rate of vitamin D testing between 2005 and 2015, secondly on prevalence of patients having a test result below 30 nmol/L and thirdly on factors influencing the test results. The manuscript is well written and there is a lot of interesting data. Globally, I think this work is very interesting and need to be published, but could be further improved mainly by accounting for osteoporosis treatments.
------------------	--

Major comments:

- -Events included calcitriol testings (1,25-dihydroxyvitamin D). I think they should be analyzed separately as they correspond to a different clinical situation compared to calcidiol (25OHD) investigation. Calcitriol testing is not suited to determine if the patient has a vitamin D deficiency (partly because of its shorter half-life compared to calcidiol, aka hours vs weeks, which turn it inappropriate to evaluate stocks of vitamin D. see Brandi, M.L. (2010). "Indications on the use of vitamin D and vitamin D metabolites in clinical phenotypes". Clinical Cases in Mineral and Bone Metabolism. 7 (3): 243–250.). Thus, include them in the logistic regression is not appropriate and include them in the rate calculus somewhat arguable. However, it could be interesting to know if the evolution of calcitriol testing followed that of calcidiol testings. It could indicate that one is used for the other, which could be worrying.
- There is a huge gap to fill regarding osteoporosis treatment (OPT) influence. It is not developed in the manuscript albeit it could explain a lot of the results (e.g why older women are more likely to be tested, but more prone to be vitamin D sufficient). Current recommendation regarding osteoporosis management is first to restore vitamin D levels before setting up OPT (which often encompass vitamin D and calcium supplementation). A lot of practionners first give vitamin D without testing and then control that vitamin D level is adequate before setting up the following step, e.g bisphosphonates initiation. It's a major protecting factor regarding vitamin D deficiency, and could be also a major confounding factor.
- The discussion state that there was no decrease of incidence rates of vitamin D testings, however there is an inflexion in the sigmoid curve since 2013 (fig 1). This should be accounted for in the discussion of recommendations efficacy.

Minor remarks:

- p.7, lines 27-28: acceptable mortality recoding or recording?
 p.9, lines 12-13: why differentiate British vs Others? I would recommend to classify into Asian, Black, White, Mixed and Others, including in tables.
- P.9, lines 39-33: I wonder regarding the adequacy of reporting a the overall incidence over 10 years (6.2), knowing that there is a huge difference between 2005 (0.5) and 2015 (23.2). I think in this case this information (the mean) is meaningless and stating the two values (2005 and 2015) is sufficient.
- P.10, 14-16: "There is a trend for a lower risk of being deficient with older age". It could be better of not speaking of trend without having had tested it formally. Moreover, regarding age and vitamin D deficiency, we observe U-shape relationship in women, which is not surprising regarding the relationship between osteoporosis treatment and aging.
- P13, 8-9: 25 nmol/L or 30 nmol/L ? Which work is commented here?
- In order to strengthen your external validity, you could also mention our work focusing a similar question (ok, I'm totally selling my work here, but I really think it's relevant. However it's up to you to include it or not https://www.nature.com/articles/s41598-017-10263-8. Really). Albeit we focused on repetition of dosing in individuals, we seen globally the same evolution than the one you observed. You could also find how French institutions reacted when faced with this brutal increase in dosing (teaser inside).

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Reviewer Name: Emre Basatemur

Institution and Country: Institute of Child Health University College London UK Please state any competing interests or state 'None declared': None declared

This manuscript describes analysis of a UK primary care research database (THIN), investigating trends in testing of vitamin D levels over time among adults in primary care and socio-demographic factors associated with rates of testing. Similar work has been previously published reporting rates of vitamin D testing and associated healthcare costs in children in primary care, using THIN (European Journal of Pediatrics, 2017;176(10):1405-1409). The current manuscript describes a similar time trend in testing to that which has been reported in children, with a marked increase in testing over the last decade.

The analysis has generally been conducted appropriately, and the paper is clear and well written. Overall the work is a useful addition to the literature. However, there are a number of points which I believe should be addressed prior to publication. These are detailed below.

Abstract:

1. The wording for the reporting of 25-OH-D deficiency prevalence in the abstract is currently misleading. In the conclusion the authors state that 'One-third of the UK population were vitamin D deficient'. However, the study is not an analysis of vitamin D levels in a random sample of the population, and the prevalence of deficiency among those tested is unlikely to be representative of that in the population as a whole. This should be re-worded to state 'One third of UK adults who had a vitamin D test performed in primary care were found to be vitamin D deficient'.

Yes, we agree with the reviewer that this statement is misleading and have altered the wording to (P2):

"One third of UK adults who had a vitamin D test performed in primary care were vitamin D deficient".

Methods:

1. Why did the authors include 1,25-dihydroxyvitamin D tests in the analysis? This is not a measure of vitamin D status nor a test for vitamin D deficiency. It is a specialist test indicated in patients with renal disease or parathyroid dysfunction. It is very different from 25-OH-D tests used to assess vitamin D status, and it should not be combined with 25-OH-D tests as a single outcome. I would recommend removing 1,25-dihydroxyvitamin D tests from the study outcome, as it is not relevant to the aims of the study (time trends in testing for vitamin D deficiency).

In light of both reviewers' comments querying the inclusion of calcitriol tests, we have removed these tests (which were only a very small number of vitamin D tests: n=1912) from the analysis and this has reduced the number of patients who were tested for vitamin D deficiency from 212,414 to 210,502. All the results have been altered accordingly, and the main findings remain unchanged.

2. I find the methods used for handling missing data for ethnicity and SES slightly confusing. For the main analysis, the authors have handled missing data by introducing a separate category indicating missingness. However, they have then not included this category when reporting the results in Table 2 (instead giving the complete results for this model in Supplementary Table 6). Table 2 should include all categories for each variable included in the model, including the one for missing. It is misleading to exclude a category for a variable when reporting the results. Table 2 should be updated to include the missing category, then Supplementary Table 6 removed.

As a sensitivity analysis, the authors performed the analysis using complete cases analysis. However, both are simple methods of dealing with missing data that are only valid if data are <imissing completely at random</i, and the results are unlikely to differ much between methods. If the authors wanted to undertake a more sophisticated sensitivity analysis, then they could go down the multiple imputation route. If not, then my personal preference would be for the authors to just present the main results using either the 'Indicator variable for missingness' method or the 'complete cases' method, and leave it at that (without reporting the other of these two simple methods as a "sensitivity analysis").

In light of the reviewer's comment we now present the results for the "Not known" category throughout all the results and text relating to "sensitivity analysis" and the corresponding supplementary tables have been removed.

3. Why did the authors choose the cut-off of <30 nmol/l to represent vitamin D deficiency? This should be explained. The National Osteoporosis Society guidelines (which they refer to) and NICE guidelines for management of vitamin D deficiency in adults both use a cut-off of <25 nmol/l to recommend treatment, so I am surprised that the authors chose <30. It would also be helpful if the authors included a comment in the discussion about the lack of consensus internationally regarding the choice of 25-OH-D threshold to indicate deficiency.

We chose a cut-off of < 30 nmol/L to indicate vitamin D deficiency based on the National Osteoporosis guideline (National Osteoporosis Society 2013), which states the following in relation to defining vitamin D deficiency (page 5):

"In agreement with the Institute of Medicine (IOM), we propose that the following vitamin D thresholds are adopted by UK practitioners in respect to bone health:

serum 25OHD < 30 nmol/L is deficient

serum 25OHD of 30-50 nmol/L may be inadequate in some people

serum 25OHD > 50 nmol/L is sufficient for almost the whole population."

We agree there is a lack of global consensus regarding the choice of a cut-point in 25(OH)D concentrations to indicate vitamin D deficiency and in light of this comment, we have added the following to the discussion (P11):

- "...and choice of a higher cut point to define vitamin D deficiency (< 30 nmol/L) in this analysis. There is a lack of global consensus regarding the choice of cut points used to define vitamin D deficiency meaning that it can often be difficult to quantify the extent of vitamin D deficiency and compare rates across studies."
- 4. Why did the authors exclude patients with 25-OH-D levels <5 or 200 in the analysis of prevalence of vitamin d deficiency among the tested individuals? Did they decide that these were biologically implausible 'outlier' values? I would argue that these values should not be treated as outliers, as they are clinically plausible. It is not unusual to see results of <5 in deficient individuals, and also high levels 200 can be seen in individuals taking high dose supplementation. Removing these test values will bias the results. Perhaps the authors chose to remove them because a result of "<5" or "200" does not represent an integer value that can be fitted into a mathematical model. In this case, I would suggest a preferable solution would be to replace these results with a value of 4 or 201 respectively rather than to ignore them. Simply excluding the extreme (but plausible) test results will lead to results that are mis-representative of the true distribution of deficiency.

To address the reviewer's comments, we have rerun the analysis, setting the value of 25(OH)D among patients with a value less than 5 nmol/L to 4 nmol/L and setting the value of patients with a value greater than 200 nmol/L to 201 nmol/L. All the results have been altered accordingly.

Results:

1. Minor point. The title for Supplementary Table 1 is missing.

The title for Supplementary Table 1 has been added in.

Conclusions:

1. The authors state that the results of this paper cannot be directly compared to the published work investigating vitamin D testing and diagnosis in children in UK primary care. However, they are perhaps unaware of a paper which has reported rates of vitamin D testing among children in UK primary care (European Journal of Pediatrics, 2017;176(10):1405-1409, see Supplementary Table 1 in this paper). This paper reports very similar time trends in vitamin D testing among children. It would be helpful if the authors compare their work to this existing paediatric literature.

We thank the reviewer for drawing our attention to this paper and now compare rates of testing in adults to that in children based on this paper in the discussion (P11):

"Others have shown a dramatic increase in the rates of testing for vitamin D among children in primary care from 2008 to 2014 and have also shown that children with Asian, black, mixed and other ethnicity were more likely to be tested compared to white children"

2. The authors hypothesise that the decrease in percentage of tested patients with deficient results over time may be due to changes in laboratory assay methods. However, I would argue that a much more likely explanation for this observation is that vitamin D testing is being performed less selectively (indicated by the large increase in testing rates over time).

To address the reviewer's comment, we have added the following into the discussion (P12):

"Therefore, an alternative explanation for the small decrease in the proportion of patients with vitamin D deficiency over time could be due to the large increase in the number of patients undergoing testing for vitamin D deficiency, suggesting that testing for vitamin D deficiency is being performed less selectively over time"

3. The cost analysis on page 13 is very misleading as only one test per individual is included, and therefore represents an underestimation of true costs. If the authors wish to make a meaningful estimate for primary care expenditure on vitamin D tests in adults, they need to include all tests (including repeat tests in the same individual). Otherwise it is better to leave this out altogether. See European Journal of Pediatrics, 2017;176(10):1405-1409 for an analysis UK primary care expenditure on vitamin D tests in children, where repeat tests within individuals were included.

In light of the reviewer's comment, we have omitted the cost analysis on page 13.

Article Summary:

1. The authors state that incidence rates for vitamin D testing in UK primary care have not previously been examined. This is not entirely accurate, as it has been examined in children (European Journal of Pediatrics, 2017;176(10):1405-1409). The authors should qualify that incidence rates for vitamin D testing in UK primary care have not previously been examined in adults.

To address this comment, we now state the following in the article summary (P14):

"...and incidence rates of testing have not been examined before in adults."

Reviewer: 2

Reviewer Name: Pascal CAILLET

Institution and Country: Nantes' University Hospital Please state any competing interests or state 'None declared': None declared

The manuscript describe the results of an open cohort analysis, focusing firstly on establishing the incidence rate of vitamin D testing between 2005 and 2015, secondly on prevalence of patients having a test result below 30 nmol/L and thirdly on factors influencing the test results.

The manuscript is well written and there is a lot of interesting data. Globally, I think this work is very interesting and need to be published, but could be further improved mainly by accounting for osteoporosis treatments.

Major comments:

- Events included calcitriol testings (1,25-dihydroxyvitamin D). I think they should be analyzed separately as they correspond to a different clinical situation compared to calcidiol (25OHD) investigation. Calcitriol testing is not suited to determine if the patient has a vitamin D deficiency (partly because of its shorter half-life compared to calcidiol, aka hours vs weeks, which turn it inappropriate to evaluate stocks of vitamin D. see Brandi, M.L. (2010). "Indications on the use of vitamin D and vitamin D metabolites in clinical phenotypes". Clinical Cases in Mineral and Bone Metabolism. 7 (3): 243–250.). Thus, include them in the logistic regression is not appropriate and include them in the rate calculus somewhat arguable. However, it could be interesting to know if the evolution of calcitriol testing followed that of calcidiol testings. It could indicate that one is used for the other, which could be worrying.

In light of both reviewers' comments querying the inclusion of calcitriol tests, we have removed these tests from the analysis and this has reduced the number of patients who were tested for vitamin D deficiency from 212,414 to 210,502. All figures in the tables and the manuscript have been altered accordingly.

- There is a huge gap to fill regarding osteoporosis treatment (OPT) influence. It is not developed in the manuscript albeit it could explain a lot of the results (e.g why older women are more likely to be tested, but more prone to be vitamin D sufficient). Current recommendation regarding osteoporosis management is first to restore vitamin D levels before setting up OPT (which often encompass vitamin D and calcium supplementation). A lot of practionners first give vitamin D without testing and then control that vitamin D level is adequate before setting up the following step, e.g bisphosphonates initiation. It's a major protecting factor regarding vitamin D deficiency, and could be also a major confounding factor.

We agree that treatment for osteoporosis could be a potential confounder in the rates of vitamin D deficiency for older women but we think it is beyond the scope of the current manuscript to investigate whether older women are having their vitamin D levels restored before beginning osteoporosis treatment. In light of the reviewer's comment, we have added in the following to the manuscript which mentions the possibility that older women may have been prescribed vitamin D as part of an osteoporosis treatment plan which may explain why they were more likely to be tested for vitamin D deficiency but less likely to be vitamin D deficient (P11-12):

"The finding of higher rates of testing for vitamin D deficiency but lower rates of deficiency in older compared with younger women could be because these women were more likely to be diagnosed with osteoporosis and have vitamin D deficiency corrected before beginning treatment. If they were taking vitamin D supplements before being tested for vitamin D deficiency then this would explain the lower rates of vitamin D deficiency. Although it should be noted that NOS does not recommend routine testing for vitamin D deficiency in patients who are co-prescribed treatment for osteoporosis and a vitamin D supplement."

- The discussion state that there was no decrease of incidence rates of vitamin D testings, however there is an inflexion in the sigmoid curve since 2013 (fig 1). This should be accounted for in the discussion of recommendations efficacy.

Yes, there does appear to be a slight inflection in the incidence curve in 2013. To address the reviewer's comment, we have added the following to the discussion (P12):

"..., although the incidence curve suggested a slight inflection point in 2013."

Minor remarks:

- p.7, lines 27-28: acceptable mortality recoding or recording?

We thank the reviewer for pointing this mistake out and have changed the text to "mortality recording".

- p.9, lines 12-13: why differentiate British vs Others? I would recommend to classify into Asian, Black, White, Mixed and Others, including in tables.

The grouping for ethnicity was based on five groups that were used in the UK National Census. The "other" group includes Chinese, Middle Eastern and Pacific ethnicity so is different from "mixed" ethnicity and other researchers using the THIN database have also categorised ethnicity in this way and we would prefer to keep the categorisation consistent. To address the reviewer's comment we have added in more detail to describe the ethnic groups included in the "other" group (P6):

"Ethnicity was categorised into five groups based on that used in the UK Census; white, black or black British, mixed, Asian or Asian British, and other (which includes Chinese, Middle Eastern and Pacific)."

- P.9, lines 39-33: I wonder regarding the adequacy of reporting a the overall incidence over 10 years (6.2), knowing that there is a huge difference between 2005 (0.5) and 2015 (23.2). I think in this case this information (the mean) is meaningless and stating the two values (2005 and 2015) is sufficient.

We have removed the average incidence for the 10 years in this section of the manuscript and now only report incidence rates for 2005 and 2015.

- P.10, 14-16: "There is a trend for a lower risk of being deficient with older age". It could be better of not speaking of trend without having had tested it formally. Moreover, regarding age and vitamin D deficiency, we observe U-shape relationship in women, which is not surprising regarding the relationship between osteoporosis treatment and aging.
- P13, 8-9: 25 nmol/L or 30 nmol/L? Which work is commented here?

The National Diet and Nutrition Survey (NDNS) used a cut-point of 25 nmol/L to define vitamin D deficiency, whereas we have chosen a cut point of 30 nmol/L based on guidance from the National Osteoporosis Society. We have altered the following in the manuscript to clarify the differences in cut points used to define deficiency (P11):

- "...and choice of a higher cut point to define vitamin D deficiency (< 30 nmol/L) in this analysis. There is a lack of global consensus regarding the choice of cut points used to define vitamin D deficiency meaning that it can often be difficult to quantify the extent of vitamin D deficiency and compare rates across studies."
- In order to strengthen your external validity, you could also mention our work focusing a similar question (ok, I'm totally selling my work here, but I really think it's relevant. However it's up to you to include it or not https://www.nature.com/articles/s41598-017-10263-8. Really). Albeit we focused on repetition of dosing in individuals, we seen globally the same evolution than the one you observed. You could also find how French institutions reacted when faced with this brutal increase in dosing (teaser inside).

Thank you for referring us to your manuscript, which is very relevant to this study. We now cite this paper in relevant sections throughout the manuscript.

VERSION 2 – REVIEW

REVIEWER	Emre Basatemur
	Institute of Child Health, University College London, UK
REVIEW RETURNED	27-Mar-2019

GENERAL COMMENTS	The authors have appropriately addressed the peer review comments for the original submission, and have made corresponding changes to the analysis and manuscript.
	I feel that the analysis has been conducted appropriately, and the manuscript is well written and clear. The work demonstrates a marked increase in testing for vitamin D deficiency among adults in UK primary care over the last decade. The work is a useful addition to the literature, and mirrors similar trends reported in a number of other countries. I recommend that the manuscript is accepted for publication.

REVIEWER	Pascal Caillet Nante's University Hospital, France
REVIEW RETURNED	13-Mar-2019

GENERAL COMMENTS	All my comments have been taken into account in an appropriate
	way. It's a great job, congratulations!